



SEQUENCE LISTING

<110> Gravel, Roy A.
Rozen, Rima
LeClerc, Daniel
Goyette, Philippe
Campeau, Eric

<120> HUMAN METHIONINE SYNTHASE: CLONING, AND
METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
CARDIOVASCULAR DISEASE, AND CANCER

<130> 50004/002005

<140> 10/607,712

<141> 2003-06-27

<150> 08/980,326

<151> 1997-11-26

<150> 60/031,964

<151> 1996-11-27

<150> 60/050,310

<151> 1997-06-20

<160> 76

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 3919

<212> DNA

<213> Homo sapiens

<220>

<221> Other

<222> (1)...(3919)

<223> Entire cloned cDNA encoding wild type methionine
synthase.

<400> 1

```
gggtcacctgt ggagagcacg tcttctctgc cgcgccctct gcgcaaggag gagactcgac 60
aacatgtcac ccgcgctcca agacctgtcg caaccggaag gtctgaagaa aaccctgcgg 120
gatgagatca atgccattct gcagaagagg attatggtgc tggatggagg gatggggacc 180
atgatccagc gggagaagct aaacgaagaa cacttccgag gtcaggaatt taaagatcat 240
gccaggccgc tgaaaggcaa caatgacatt ttaagtataa ctcagcctga tgtcatttac 300
caaatccata aggaatactt gctggctggg gcagatatca ttgaaacaaa tacttttagc 360
agcactagta ttgcccaagc tgactatggc cttgaacact tggcctaccg gatgaacatg 420
tgctctgcag gagtggccag aaaagctgcc gaggaggtaa ctctccagac aggaattaag 480
aggtttgttg caggggctct ggggtccgact aataagacac tctctgtgtc cccatctgtg 540
gaaaggccgg attataggaa catcacattt gatgagcttg ttgaagcata ccaagagcag 600
gccaaaggac ttctggatgg cggggttgat atcttactca ttgaaactat ttttgatact 660
gccaatgcca aggcagcctt gtttgcactc caaaatcttt ttgaggagaa atatgctccc 720
cggcctatct ttatttcagg gacgatcggt gataaaagtg ggcggactct ttccggacag 780
acaggagagg gatttgcacat cagcgtgtct catggagaac cactctgcat tggattaaat 840
tgtgctttgg gtgcagctga gatgagacct tttattgaaa taattggaaa atgtacaaca 900
```

gcctatgtcc	tctgttatcc	caatgcaggt	cttcccaaca	ccttttggtga	ctatgatgaa	960
acgccttcta	tgatggccaa	gcacctaaag	gatttttgcta	tggatggctt	ggtcaatata	1020
gttggaggat	gctgtgggtc	aacaccagat	catatcaggg	aaattgctga	agctgtgaaa	1080
aattgtaagc	ctagagttcc	acctgccact	gcttttgaag	gacatatgtt	actgtctggt	1140
ctagagccct	tcaggattgg	accgtacacc	aactttgtta	acattggaga	gcgctgtaat	1200
gttgcaggat	caaggaagtt	tgctaaactc	atcatggcag	gaaactatga	agaagccttg	1260
tgtgttgcca	aagtgcaggt	ggaaatggga	gccaggtgt	tggatgtcaa	catggatgat	1320
ggcatgctag	atggtccaag	tgcaatgacc	agattttgca	acttaattgc	ttccgagcca	1380
gacatcgcaa	aggtaccttt	gtgcatcgac	tcctccaatt	ttgctgtgat	tgaagctggg	1440
ttaaagtgtc	gccaaaggaa	gtgcattgtc	aatagcatta	gtctgaagga	aggagaggac	1500
gacttcttgg	agaaggccag	gaagattaaa	aagtatggag	ctgctatggt	ggtcatggct	1560
tttgatgaag	aaggacaggc	aacagaaaca	gacacaaaaa	tcagagtgtg	cacccgggcc	1620
taccatctgc	ttgtgaaaaa	actgggcttt	aatccaaatg	acattatttt	tgaccctaata	1680
atcctaacca	ttgggactgg	aatggaggaa	cacaacttgt	atgccattaa	ttttatccat	1740
gcaacaaaag	tcattaaaga	aacattacct	ggagccagaa	taagtggagg	tcctttccaac	1800
ttgtccttct	ccttccgagg	aatggaagcc	attcgagaag	caatgcatgg	ggttttccct	1860
taccatgcaa	tcaagtctgg	catggacatg	gagatagtga	atgctggaaa	cctccctgtg	1920
tatgatgata	tccataagga	acttctgcag	ctctgtgaag	atctcatctg	gaataaagac	1980
cctgaggcca	ctgagaagct	cttacgttat	gccagactc	aaggcacagg	agggaagaaa	2040
gtcattcaga	ctgatgagt	gagaaatggc	cctgtcgaag	aacgccttga	gtatgccctt	2100
gtgaagggca	ttgaaaaaca	tattattgag	gatactgagg	aagccagggt	aaacaaaaaa	2160
aaatatcccc	gacctctcaa	tataattgaa	ggacccctga	tgaatggaat	gaaaattgtt	2220
ggtgatcttt	ttggagctgg	aaaaatgttt	ctacctcagg	ttataaagtc	agcccggtt	2280
atgaagaagg	ctgttggcca	ccttatccct	ttcatggaaa	aagaaagaga	agaaaccaga	2340
gtgcttaacg	gcacagttag	agaagaggac	ccttaccagg	gcaccatcgt	gctggccact	2400
gttaaaaggc	acgtgcacga	cataggcaag	aacatagttg	gagtagtcct	tggctgcaat	2460
aatttccgag	ttattgattt	aggagtcatg	actccatgtg	ataagatact	gaaagctgct	2520
cttgaccaca	aagcagatat	aattggcctg	tcaggactca	tcactccttc	cctggatgaa	2580
atgatttttg	ttgccaaagg	aatggagaga	ttagctataa	ggattccatt	gttgattgga	2640
ggagcaacca	cttcaaaaac	ccacacagca	gttaaaatag	ctccgagata	cagtgcacct	2700
gtaatccatg	tcctggacgc	gtccaagagt	gtggtggtgt	gttcccagct	gttagatgaa	2760
aatctaaagg	atgaatactt	tgaggaaatc	atggaagaat	atgaagatat	tagacaggac	2820
cattatgagt	ctctcaagga	gaggagatac	ttacccttaa	gtcaagccag	aaaaagtgtg	2880
ttccaaatgg	attggctgtc	tgaacctcac	ccagtgaagc	ccacgtttat	tgggaccag	2940
gtctttgaag	actatgacct	gcagaagctg	gtggactaca	ttgactggaa	gcctttcttt	3000
gatgtctggc	agctccgggg	caagtaccgc	aatcgaggct	tccccaagat	atttaacgac	3060
aaaacagtag	gtggagaggc	caggaaggtc	tacgatgatg	cccacaatat	gctgaacaca	3120
ctgattagtc	aaaagaaact	ccgggcccgc	ggtgtggttg	ggttctggcc	agcacagagt	3180
atccaagacg	acattcacct	gtacgcagag	gctgctgtgc	cccaggctgc	agagcccata	3240
gccactttct	atgggttaag	gcaacaggct	gagaaggact	ctgccagcac	ggagccatac	3300
tactgcctct	cagacttcat	cgctcccttg	cattctggca	tcctgtgacta	cctgggcttg	3360
tttgccgttg	cctgcttttg	ggtagaagag	ctgagcaagg	cctatgagga	tgatggtgac	3420
gactacagca	gcatcatggt	caaggcgtcg	ggggaccggc	tggcagaggc	ctttgcagaa	3480
gagctccatg	aaagagttcg	ccgagaactg	tgggcctact	gtggcagtga	gcagctggac	3540
gtcgcagacc	tgcgaagggt	gcggtacaag	ggcatccgcc	cggtccttg	ctaccccagc	3600
cagcccagacc	acaccgagaa	gctcaccatg	tggagactcg	cagacatcga	gcagtctaca	3660
ggcattaggt	taacagaatc	attagcaatg	gcacctgctt	cagcagtctc	aggcctctac	3720
ttctccaatt	tgaagtccaa	atattttgct	gtggggaaga	tttccaagga	tcaggttgag	3780
gattatgcat	tgaggaagaa	catatctgtg	gctgaggttg	agaaatggct	tggaccatt	3840
ttgggatatg	atacagacta	actttttttt	ttttttttgc	cttttttatc	ttgatgatcc	3900
tcaaggaaat	acaacctag					3919

<210> 2
 <211> 1265
 <212> PRT
 <213> Homo sapiens

<220>
 <221> VARIANT

<222> (1)...(1265)

<223> Wild type methionine synthase polypeptide.

<400> 2

Met	Ser	Pro	Ala	Leu	Gln	Asp	Leu	Ser	Gln	Pro	Glu	Gly	Leu	Lys	Lys
1				5					10					15	
Thr	Leu	Arg	Asp	Glu	Ile	Asn	Ala	Ile	Leu	Gln	Lys	Arg	Ile	Met	Val
			20					25					30		
Leu	Asp	Gly	Gly	Met	Gly	Thr	Met	Ile	Gln	Arg	Glu	Lys	Leu	Asn	Glu
		35					40					45			
Glu	His	Phe	Arg	Gly	Gln	Glu	Phe	Lys	Asp	His	Ala	Arg	Pro	Leu	Lys
	50					55					60				
Gly	Asn	Asn	Asp	Ile	Leu	Ser	Ile	Thr	Gln	Pro	Asp	Val	Ile	Tyr	Gln
65					70					75					80
Ile	His	Lys	Glu	Tyr	Leu	Leu	Ala	Gly	Ala	Asp	Ile	Ile	Glu	Thr	Asn
			85					90						95	
Thr	Phe	Ser	Ser	Thr	Ser	Ile	Ala	Gln	Ala	Asp	Tyr	Gly	Leu	Glu	His
			100					105					110		
Leu	Ala	Tyr	Arg	Met	Asn	Met	Cys	Ser	Ala	Gly	Val	Ala	Arg	Lys	Ala
		115					120					125			
Ala	Glu	Glu	Val	Thr	Leu	Gln	Thr	Gly	Ile	Lys	Arg	Phe	Val	Ala	Gly
	130					135					140				
Ala	Leu	Gly	Pro	Thr	Asn	Lys	Thr	Leu	Ser	Val	Ser	Pro	Ser	Val	Glu
145					150					155					160
Arg	Pro	Asp	Tyr	Arg	Asn	Ile	Thr	Phe	Asp	Glu	Leu	Val	Glu	Ala	Tyr
			165					170						175	
Gln	Glu	Gln	Ala	Lys	Gly	Leu	Leu	Asp	Gly	Gly	Val	Asp	Ile	Leu	Leu
			180					185					190		
Ile	Glu	Thr	Ile	Phe	Asp	Thr	Ala	Asn	Ala	Lys	Ala	Ala	Leu	Phe	Ala
	195					200					205				
Leu	Gln	Asn	Leu	Phe	Glu	Glu	Lys	Tyr	Ala	Pro	Arg	Pro	Ile	Phe	Ile
	210					215					220				
Ser	Gly	Thr	Ile	Val	Asp	Lys	Ser	Gly	Arg	Thr	Leu	Ser	Gly	Gln	Thr
225					230					235					240
Gly	Glu	Gly	Phe	Val	Ile	Ser	Val	Ser	His	Gly	Glu	Pro	Leu	Cys	Ile
			245						250					255	
Gly	Leu	Asn	Cys	Ala	Leu	Gly	Ala	Ala	Glu	Met	Arg	Pro	Phe	Ile	Glu
		260						265					270		
Ile	Ile	Gly	Lys	Cys	Thr	Thr	Ala	Tyr	Val	Leu	Cys	Tyr	Pro	Asn	Ala
	275						280					285			
Gly	Leu	Pro	Asn	Thr	Phe	Gly	Asp	Tyr	Asp	Glu	Thr	Pro	Ser	Met	Met
	290					295					300				
Ala	Lys	His	Leu	Lys	Asp	Phe	Ala	Met	Asp	Gly	Leu	Val	Asn	Ile	Val
305					310					315					320
Gly	Gly	Cys	Cys	Gly	Ser	Thr	Pro	Asp	His	Ile	Arg	Glu	Ile	Ala	Glu
			325						330					335	
Ala	Val	Lys	Asn	Cys	Lys	Pro	Arg	Val	Pro	Pro	Ala	Thr	Ala	Phe	Glu
		340						345					350		
Gly	His	Met	Leu	Leu	Ser	Gly	Leu	Glu	Pro	Phe	Arg	Ile	Gly	Pro	Tyr
	355						360					365			
Thr	Asn	Phe	Val	Asn	Ile	Gly	Glu	Arg	Cys	Asn	Val	Ala	Gly	Ser	Arg
	370					375					380				
Lys	Phe	Ala	Lys	Leu	Ile	Met	Ala	Gly	Asn	Tyr	Glu	Glu	Ala	Leu	Cys
385					390					395					400
Val	Ala	Lys	Val	Gln	Val	Glu	Met	Gly	Ala	Gln	Val	Leu	Asp	Val	Asn
			405						410					415	
Met	Asp	Asp	Gly	Met	Leu	Asp	Gly	Pro	Ser	Ala	Met	Thr	Arg	Phe	Cys
			420					425					430		
Asn	Leu	Ile	Ala	Ser	Glu	Pro	Asp	Ile	Ala	Lys	Val	Pro	Leu	Cys	Ile

Tyr Glu Asp Ile Arg Gln Asp His Tyr Glu Ser Leu Lys Glu Arg Arg
 915 920 925
 Tyr Leu Pro Leu Ser Gln Ala Arg Lys Ser Gly Phe Gln Met Asp Trp
 930 935 940
 Leu Ser Glu Pro His Pro Val Lys Pro Thr Phe Ile Gly Thr Gln Val
 945 950 955 960
 Phe Glu Asp Tyr Asp Leu Gln Lys Leu Val Asp Tyr Ile Asp Trp Lys
 965 970 975
 Pro Phe Phe Asp Val Trp Gln Leu Arg Gly Lys Tyr Pro Asn Arg Gly
 980 985 990
 Phe Pro Lys Ile Phe Asn Asp Lys Thr Val Gly Gly Glu Ala Arg Lys
 995 1000 1005
 Val Tyr Asp Asp Ala His Asn Met Leu Asn Thr Leu Ile Ser Gln Lys
 1010 1015 1020
 Lys Leu Arg Ala Arg Gly Val Val Gly Phe Trp Pro Ala Gln Ser Ile
 1025 1030 1035 1040
 Gln Asp Asp Ile His Leu Tyr Ala Glu Ala Val Pro Gln Ala Ala
 1045 1050 1055
 Glu Pro Ile Ala Thr Phe Tyr Gly Leu Arg Gln Gln Ala Glu Lys Asp
 1060 1065 1070
 Ser Ala Ser Thr Glu Pro Tyr Tyr Cys Leu Ser Asp Phe Ile Ala Pro
 1075 1080 1085
 Leu His Ser Gly Ile Arg Asp Tyr Leu Gly Leu Phe Ala Val Ala Cys
 1090 1095 1100
 Phe Gly Val Glu Glu Leu Ser Lys Ala Tyr Glu Asp Asp Gly Asp Asp
 1105 1110 1115 1120
 Tyr Ser Ser Ile Met Val Lys Ala Leu Gly Asp Arg Leu Ala Glu Ala
 1125 1130 1135
 Phe Ala Glu Glu Leu His Glu Arg Val Arg Arg Glu Leu Trp Ala Tyr
 1140 1145 1150
 Cys Gly Ser Glu Gln Leu Asp Val Ala Asp Leu Arg Arg Leu Arg Tyr
 1155 1160 1165
 Lys Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ser Gln Pro Asp His Thr
 1170 1175 1180
 Glu Lys Leu Thr Met Trp Arg Leu Ala Asp Ile Glu Gln Ser Thr Gly
 1185 1190 1195 1200
 Ile Arg Leu Thr Glu Ser Leu Ala Met Ala Pro Ala Ser Ala Val Ser
 1205 1210 1215
 Gly Leu Tyr Phe Ser Asn Leu Lys Ser Lys Tyr Phe Ala Val Gly Lys
 1220 1225 1230
 Ile Ser Lys Asp Gln Val Glu Asp Tyr Ala Leu Arg Lys Asn Ile Ser
 1235 1240 1245
 Val Ala Glu Val Glu Lys Trp Leu Gly Pro Ile Leu Gly Tyr Asp Thr
 1250 1255 1260
 Asp
 1265

<210> 3
 <211> 9
 <212> PRT
 <213> Escherichia coli

<400> 3
 Asp Gly Gly Met Gly Thr Met Ile Gln
 1 5

<210> 4

<211> 9
<212> PRT
<213> *Cyanobacterium synechocystis*

<400> 4
Asp Gly Ala Met Gly Thr Asn Leu Gln
1 5

<210> 5
<211> 9
<212> PRT
<213> *Mycobacterium leprae*

<400> 5
Asp Gly Ala Met Gly Thr Gln Leu Gln
1 5

<210> 6
<211> 9
<212> PRT
<213> *Hemophilus influenzae*

<400> 6
Asp Gly Ala Met Gly Thr Met Ile Gln
1 5

<210> 7
<211> 9
<212> PRT
<213> *Caenorhabditis elegans*

<400> 7
Asp Gly Ala Met Gly Thr Met Ile Gln
1 5

<210> 8
<211> 9
<212> PRT
<213> *Homo sapiens*

<400> 8
Asp Gly Gly Met Gly Thr Met Ile Gln
1 5

<210> 9
<211> 13
<212> PRT
<213> *Escherichia coli*

<400> 9
Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn
1 5 10

<210> 10
 <211> 13
 <212> PRT
 <213> *Cyanobacterium synechocystis*

 <400> 10
 Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn
 1 5 10

 <210> 11
 <211> 13
 <212> PRT
 <213> *Mycobacterium leprae*

 <400> 11
 Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn
 1 5 10

 <210> 12
 <211> 13
 <212> PRT
 <213> *Hemophilus influenzae*

 <400> 12
 Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn
 1 5 10

 <210> 13
 <211> 13
 <212> PRT
 <213> *Caenorhabditis elegans*

 <400> 13
 Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn
 1 5 10

 <210> 14
 <211> 13
 <212> PRT
 <213> *Homo sapiens*

 <400> 14
 Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn
 1 5 10

 <210> 15
 <211> 10
 <212> PRT
 <213> *Escherichia coli*

 <400> 15
 Leu Ala Glu Ala Phe Ala Glu Tyr Leu His
 1 5 10

<210> 16
<211> 10
<212> PRT
<213> Cyanobacterium synechocystis

<400> 16
Met Ala Glu Ala Leu Ala Glu Trp Thr His
1 5 10

<210> 17
<211> 10
<212> PRT
<213> Mycobacterium leprae

<400> 17
Leu Thr Glu Ala Leu Ala Glu Tyr Trp His
1 5 10

<210> 18
<211> 10
<212> PRT
<213> Hemophilus influenzae

<400> 18
Leu Ala Glu Ala Met Ala Glu Tyr Leu His
1 5 10

<210> 19
<211> 10
<212> PRT
<213> Caenorrrhabditis elegans

<400> 19
Leu Ala Glu Ala Tyr Ala Glu Tyr Leu His
1 5 10

<210> 20
<211> 10
<212> PRT
<213> Homo sapiens

<400> 20
Leu Ala Glu Ala Phe Ala Glu Glu Leu His
1 5 10

<210> 21
<211> 11
<212> PRT
<213> Escherichia coli

<400> 21
Gly Gly Cys Cys Gly Thr Thr Pro Gln His Ile
1 5 10

<210> 22
 <211> 11
 <212> PRT
 <213> Cyanobacterium synechocystis

<400> 22
 Gly Gly Cys Cys Gly Thr Arg Pro Asp His Ile
 1 5 10

<210> 23
 <211> 11
 <212> PRT
 <213> Mycobacterium leprae

<400> 23
 Gly Gly Cys Cys Gly Thr Thr Pro Asp His Ile
 1 5 10

<210> 24
 <211> 11
 <212> PRT
 <213> Caenorhabditis elegans

<400> 24
 Gly Gly Cys Cys Gly Thr Thr Pro Asp His Ile
 1 5 10

<210> 25
 <211> 11
 <212> PRT
 <213> Homo sapiens

<400> 25
 Gly Gly Cys Cys Gly Ser Thr Pro Asp His Ile
 1 5 10

<210> 26
 <211> 26
 <212> DNA
 <213> Homo sapiens

<220>
 <221> variation
 <222> (1)...(26)
 <223> n is a, t, g, or c; h is a, c, or t; d is a, g, or
 t; and r is a or g;

<400> 26
 gayggngcna tgggnacnat gathca

26

<210> 27
 <211> 29
 <212> DNA
 <213> Homo sapiens

<220>
 <221> variation
 <222> (1)...(29)
 <223> n is a, t, g, or c; h is a, c, or t; d is a, g, or
 t; and r is a or g;

 <400> 27
 gcnacngtna arggngaygt ncaygayat 29

 <210> 28
 <211> 30
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> variation
 <222> (1)...(30)
 <223> n is a, t, g, or c; h is a, c, or t; d is a, g, or
 t; and r is a or g;

 <400> 28
 rtttyttnced atrtctrtgna crtenccytt 30

 <210> 29
 <211> 27
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> variation
 <222> (1)...(27)
 <223> n is a, t, g, or c; h is a, c, or t; d is a, g, or
 t; and r is a or g;

 <400> 29
 rtgnagrtay tcngcraang cytcngc 27

 <210> 30
 <211> 32
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> variation
 <222> (1)...(32)
 <223> n is a, t, g, or c; h is a, c, or t; d is a, g, or
 t; and r is a or g;

 <400> 30
 atrtgrtcng gngtngtncc rcarcancn cc 32

 <210> 31
 <211> 32
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> variation
 <222> (1)...(32)

<223> n is a, t, g, or c; h is a, c, or t; d is a, g, or
t; and r is a or g;

<400> 31
ggnggntgyt gyggnacnac nccngaycay at 32

<210> 32
<211> 24
<212> DNA
<213> Mus musculus

<400> 32
gtctgtgtca tagcccagaa tggg 24

<210> 33
<211> 24
<212> DNA
<213> Mus musculus

<400> 33
tcagtctgtg tcatagccca gaat 24

<210> 34
<211> 24
<212> DNA
<213> Homo sapiens

<400> 34
gaactagaag acagaaattc tcta 24

<210> 35
<211> 26
<212> DNA
<213> Homo sapiens

<400> 35
ttccgaggtc aggaatttaa agatca 26

<210> 36
<211> 24
<212> DNA
<213> Homo sapiens

<400> 36
gtgttcttcg tttagcttct cccg 24

<210> 37
<211> 24
<212> DNA
<213> Homo sapiens

<400> 37
ccccagccag caagtattcc ttat 24

<210> 38
<211> 24
<212> DNA
<213> Homo sapiens

<400> 38 ctaggttgta tticcttgag gatc	24
<210> 39 <211> 25 <212> DNA <213> Homo sapiens	
<400> 39 ggagctggaa aaatgtttct acctc	25
<210> 40 <211> 24 <212> DNA <213> Homo sapiens	
<400> 40 acaggagggga agaaagtcac tcag	24
<210> 41 <211> 25 <212> DNA <213> Homo sapiens	
<400> 41 ccttcaatta tattgagagg tcggg	25
<210> 42 <211> 24 <212> DNA <213> Homo sapiens	
<400> 42 caacccgaag gtctgaagaa aacc	24
<210> 43 <211> 21 <212> DNA <213> Homo sapiens	
<400> 43 cccgcgtcc aagacctgtc g	21
<210> 44 <211> 21 <212> DNA <213> Homo sapiens	
<400> 44 cgacaggtct tggagcgagg g	21
<210> 45 <211> 28 <212> DNA <213> Homo sapiens	
<400> 45 ggagtcacga ctccataatc aataactc	28

<210> 46
 <211> 23
 <212> DNA
 <213> Homo sapiens

 <400> 46
 gacgactaca gcagcatcat ggt 23

 <210> 47
 <211> 22
 <212> DNA
 <213> Homo sapiens

 <400> 47
 aaaaatcatt tcattccagg aa 22

 <210> 48
 <211> 26
 <212> DNA
 <213> Homo sapiens

 <400> 48
 ataggcaaga acatagttgg agtagt 26

 <210> 49
 <211> 25
 <212> DNA
 <213> Homo sapiens

 <400> 49
 tttcatctaa cagctgggaa cacac 25

 <210> 50
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 50
 tgccctctcag acttcatcgc tccc 24

 <210> 51
 <211> 21
 <212> DNA
 <213> Homo sapiens

 <400> 51
 tgcagcctgg ggcacagcag c 21

 <210> 52
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 52
 atggattggc tgtctgaacc tcac 24

 <210> 53
 <211> 26
 <212> DNA

<213> Homo sapiens
 <400> 53
 catggaagaa tatgaagata ttagac 26
 <210> 54
 <211> 24
 <212> DNA
 <213> Homo sapiens
 <400> 54
 accatcatcc tcataggcct tgct 24
 <210> 55
 <211> 23
 <212> DNA
 <213> Homo sapiens
 <400> 55
 cagacctgcg aagggtgcgg tac 23
 <210> 56
 <211> 24
 <212> DNA
 <213> Homo sapiens
 <400> 56
 gaagtgggtg ctctccaat caac 24
 <210> 57
 <211> 26
 <212> DNA
 <213> Homo sapiens
 <400> 57
 gagcagcttt cagtatctta tcacat 26
 <210> 58
 <211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 58
 acaagttgtg ttctccatt ccagt 25
 <210> 59
 <211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 59
 agagcgctgt aatggtgcag gatca 25
 <210> 60
 <211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 60

tgtttttcaa tgcccttcac aaggg 25

<210> 61
<211> 25
<212> DNA
<213> Homo sapiens

<400> 61
taaaaagtat ggagctgcta tgggtg 25

<210> 62
<211> 25
<212> DNA
<213> Homo sapiens

<400> 62
gaccagacag taacatatgt ccttc 25

<210> 63
<211> 26
<212> DNA
<213> Homo sapiens

<400> 63
acattacagc gctctccaat gttaac 26

<210> 64
<211> 24
<212> DNA
<213> Homo sapiens

<400> 64
tgaggttgag aaatggcttg gacc 24

<210> 65
<211> 25
<212> DNA
<213> Homo sapiens

<400> 65
gccacagata tgttcttcct caatg 25

<210> 66
<211> 24
<212> DNA
<213> Homo sapiens

<400> 66
tgtggagagc acgtcttctc tgcc 24

<210> 67
<211> 50
<212> PRT
<213> Homo sapiens

<400> 67
Leu Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val
1 5 10 15
Gly Val Val Leu Gly Cys Asn Asn Phe Arg Val Ile Asp Leu Gly Val

20 25 30
 Met Thr Pro Cys Asp Lys Ile Leu Lys Ala Ala Leu Asp His Lys Ala
 35 40 45
 Asp Ile
 50

<210> 68
 <211> 50
 <212> PRT
 <213> Mus musculus

<400> 68
 Leu Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val
 1 5 10 15
 Gly Val Val Leu Ala Cys Asn Asn Phe Arg Val Ile Asp Leu Gly Val
 20 25 30
 Met Thr Pro Cys Asp Lys Ile Leu Gln Ala Ala Leu Asp His Lys Ala
 35 40 45
 Asp Ile
 50

<210> 69
 <211> 50
 <212> PRT
 <213> Cyanobacterium synechocystis

<400> 69
 Ile Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Leu Val
 1 5 10 15
 Asp Ile Ile Leu Ser Asn Asn Gly Tyr Arg Val Val Asn Leu Gly Ile
 20 25 30
 Lys Gln Pro Val Glu Asn Ile Ile Glu Ala Tyr Lys Lys His Arg Pro
 35 40 45
 Asp Cys
 50

<210> 70
 <211> 50
 <212> PRT
 <213> Mycobacterium leprae

<400> 70
 Leu Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Leu Val
 1 5 10 15
 Asp Ile Ile Leu Ser Asn Asn Gly Tyr Glu Val Val Asn Leu Gly Ile
 20 25 30
 Lys Gln Pro Ile Thr Asn Ile Leu Glu Val Ala Glu Asp Lys Ser Ala
 35 40 45
 Asp Val
 50

<210> 71
 <211> 50
 <212> PRT
 <213> Caenorhabditis elegans

<400> 71
 Ile Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val
 1 5 10 15
 Ser Val Val Leu Gly Cys Asn Asn Phe Lys Val Val Asp Leu Gly Val
 20 25 30
 Met Thr Pro Cys Glu Asn Ile Ile Lys Ala Ala Ile Glu Glu Lys Ala
 35 40 45
 Asp Phe
 50

<210> 72
 <211> 50
 <212> PRT
 <213> Hemophilus influenzae

<400> 72
 Ile Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val
 1 5 10 15
 Ser Val Val Met Gln Cys Asn Asn Phe Glu Val Ile Asp Leu Gly Val
 20 25 30
 Met Val Pro Ala Asp Lys Ile Ile Gln Thr Ala Ile Asn Gln Lys Thr
 35 40 45
 Asp Ile
 50

<210> 73
 <211> 50
 <212> PRT
 <213> Escherichia coli

<400> 73
 Ile Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val
 1 5 10 15
 Gly Val Val Leu Gln Cys Asn Asn Tyr Glu Ile Val Asp Leu Gly Val
 20 25 30
 Met Val Pro Ala Glu Lys Ile Leu Arg Thr Ala Lys Glu Val Asn Ala
 35 40 45
 Asp Leu
 50

<210> 74
 <211> 1265
 <212> PRT
 <213> Homo sapiens

<220>
 <221> VARIANT
 <222> (1)...(1265)
 <223> Xaa at position 881 is either Ile or no amino
 acid; Xaa at position 919 is either Asp or Gly;
 Xaa at position 920 is either His or Asp.

<400> 74
 Met Ser Pro Ala Leu Gln Asp Leu Ser Gln Pro Glu Gly Leu Lys Lys
 1 5 10 15
 Thr Leu Arg Asp Glu Ile Asn Ala Ile Leu Gln Lys Arg Ile Met Val

			20					25				30					
Leu	Asp	Gly	Gly	Met	Gly	Thr	Met	Ile	Gln	Arg	Glu	Lys	Leu	Asn	Glu		
		35					40					45					
Glu	His	Phe	Arg	Gly	Gln	Glu	Phe	Lys	Asp	His	Ala	Arg	Pro	Leu	Lys		
	50					55					60						
Gly	Asn	Asn	Asp	Ile	Leu	Ser	Ile	Thr	Gln	Pro	Asp	Val	Ile	Tyr	Gln		
65					70					75					80		
Ile	His	Lys	Glu	Tyr	Leu	Leu	Ala	Gly	Ala	Asp	Ile	Ile	Glu	Thr	Asn		
				85					90					95			
Thr	Phe	Ser	Ser	Thr	Ser	Ile	Ala	Gln	Ala	Asp	Tyr	Gly	Leu	Glu	His		
			100					105					110				
Leu	Ala	Tyr	Arg	Met	Asn	Met	Cys	Ser	Ala	Gly	Val	Ala	Arg	Lys	Ala		
	115						120					125					
Ala	Glu	Glu	Val	Thr	Leu	Gln	Thr	Gly	Ile	Lys	Arg	Phe	Val	Ala	Gly		
	130					135					140						
Ala	Leu	Gly	Pro	Thr	Asn	Lys	Thr	Leu	Ser	Val	Ser	Pro	Ser	Val	Glu		
145					150					155					160		
Arg	Pro	Asp	Tyr	Arg	Asn	Ile	Thr	Phe	Asp	Glu	Leu	Val	Glu	Ala	Tyr		
				165					170					175			
Gln	Glu	Gln	Ala	Lys	Gly	Leu	Leu	Asp	Gly	Gly	Val	Asp	Ile	Leu	Leu		
			180					185					190				
Ile	Glu	Thr	Ile	Phe	Asp	Thr	Ala	Asn	Ala	Lys	Ala	Ala	Leu	Phe	Ala		
	195						200					205					
Leu	Gln	Asn	Leu	Phe	Glu	Glu	Lys	Tyr	Ala	Pro	Arg	Pro	Ile	Phe	Ile		
	210					215					220						
Ser	Gly	Thr	Ile	Val	Asp	Lys	Ser	Gly	Arg	Thr	Leu	Ser	Gly	Gln	Thr		
225					230				235						240		
Gly	Glu	Gly	Phe	Val	Ile	Ser	Val	Ser	His	Gly	Glu	Pro	Leu	Cys	Ile		
				245					250					255			
Gly	Leu	Asn	Cys	Ala	Leu	Gly	Ala	Ala	Glu	Met	Arg	Pro	Phe	Ile	Glu		
			260					265					270				
Ile	Ile	Gly	Lys	Cys	Thr	Thr	Ala	Tyr	Val	Leu	Cys	Tyr	Pro	Asn	Ala		
	275						280					285					
Gly	Leu	Pro	Asn	Thr	Phe	Gly	Asp	Tyr	Asp	Glu	Thr	Pro	Ser	Met	Met		
	290					295					300						
Ala	Lys	His	Leu	Lys	Asp	Phe	Ala	Met	Asp	Gly	Leu	Val	Asn	Ile	Val		
305					310					315					320		
Gly	Gly	Cys	Cys	Gly	Ser	Thr	Pro	Asp	His	Ile	Arg	Glu	Ile	Ala	Glu		
				325					330					335			
Ala	Val	Lys	Asn	Cys	Lys	Pro	Arg	Val	Pro	Pro	Ala	Thr	Ala	Phe	Glu		
			340					345					350				
Gly	His	Met	Leu	Leu	Ser	Gly	Leu	Glu	Pro	Phe	Arg	Ile	Gly	Pro	Tyr		
	355						360					365					
Thr	Asn	Phe	Val	Asn	Ile	Gly	Glu	Arg	Cys	Asn	Val	Ala	Gly	Ser	Arg		
	370					375					380						
Lys	Phe	Ala	Lys	Leu	Ile	Met	Ala	Gly	Asn	Tyr	Glu	Glu	Ala	Leu	Cys		
385					390					395					400		
Val	Ala	Lys	Val	Gln	Val	Glu	Met	Gly	Ala	Gln	Val	Leu	Asp	Val	Asn		
				405					410					415			
Met	Asp	Asp	Gly	Met	Leu	Asp	Gly	Pro	Ser	Ala	Met	Thr	Arg	Phe	Cys		
			420					425					430				
Asn	Leu	Ile	Ala	Ser	Glu	Pro	Asp	Ile	Ala	Lys	Val	Pro	Leu	Cys	Ile		
			435				440					445					
Asp	Ser	Ser	Asn	Phe	Ala	Val	Ile	Glu	Ala	Gly	Leu	Lys	Cys	Cys	Gln		
	450					455					460						
Gly	Lys	Cys	Ile	Val	Asn	Ser	Ile	Ser	Leu	Lys	Glu	Gly	Glu	Asp	Asp		
465					470					475					480		
Phe	Leu	Glu	Lys	Ala	Arg	Lys	Ile	Lys	Lys	Tyr	Gly	Ala	Ala	Met	Val		
				485					490					495			

Val	Met	Ala	Phe 500	Asp	Glu	Glu	Gly	Gln 505	Ala	Thr	Glu	Thr	Asp 510	Thr	Lys
Ile	Arg	Val 515	Cys	Thr	Arg	Ala	Tyr 520	His	Leu	Leu	Val	Lys 525	Lys	Leu	Gly
Phe	Asn 530	Pro	Asn	Asp	Ile	Ile 535	Phe	Asp	Pro	Asn	Ile 540	Leu	Thr	Ile	Gly
Thr 545	Gly	Met	Glu	Glu	His 550	Asn	Leu	Tyr	Ala	Ile 555	Asn	Phe	Ile	His	Ala 560
Thr	Lys	Val	Ile	Lys 565	Glu	Thr	Leu	Pro	Gly 570	Ala	Arg	Ile	Ser	Gly 575	Gly
Leu	Ser	Asn 580	Leu	Ser	Phe	Ser	Phe	Arg 585	Gly	Met	Glu	Ala	Ile	Arg	Glu
Ala	Met	His 595	Gly	Val	Phe	Leu	Tyr 600	His	Ala	Ile	Lys	Ser 605	Gly	Met	Asp
Met	Glu 610	Ile	Val	Asn	Ala	Gly 615	Asn	Leu	Pro	Val	Tyr 620	Asp	Asp	Ile	His
Lys 625	Glu	Leu	Leu	Gln	Leu 630	Cys	Glu	Asp	Leu	Ile 635	Trp	Asn	Lys	Asp	Pro 640
Glu	Ala	Thr	Glu	Lys 645	Leu	Leu	Arg	Tyr	Ala 650	Gln	Thr	Gln	Gly	Thr 655	Gly
Gly	Lys	Lys	Val 660	Ile	Gln	Thr	Asp	Glu 665	Trp	Arg	Asn	Gly	Pro 670	Val	Glu
Glu	Arg	Leu 675	Glu	Tyr	Ala	Leu	Val 680	Lys	Gly	Ile	Glu	Lys 685	His	Ile	Ile
Glu	Asp 690	Thr	Glu	Glu	Ala	Arg 695	Leu	Asn	Gln	Lys	Lys 700	Tyr	Pro	Arg	Pro
Leu 705	Asn	Ile	Ile	Glu	Gly 710	Pro	Leu	Met	Asn	Gly 715	Met	Lys	Ile	Val	Gly 720
Asp	Leu	Phe	Gly	Ala 725	Gly	Lys	Met	Phe	Leu 730	Pro	Gln	Val	Ile	Lys 735	Ser
Ala	Arg	Val 740	Met	Lys	Lys	Ala	Val	Gly 745	His	Leu	Ile	Pro	Phe 750	Met	Glu
Lys	Glu	Arg 755	Glu	Glu	Thr	Arg	Val 760	Leu	Asn	Gly	Thr	Val 765	Glu	Glu	Glu
Asp	Pro 770	Tyr	Gln	Gly	Thr	Ile 775	Val	Leu	Ala	Thr	Val 780	Lys	Gly	Asp	Val
His 785	Asp	Ile	Gly	Lys	Asn 790	Ile	Val	Gly	Val	Val 795	Leu	Gly	Cys	Asn 800	Asn
Phe	Arg	Val	Ile	Asp 805	Leu	Gly	Val	Met	Thr 810	Pro	Cys	Asp	Lys	Ile 815	Leu
Lys	Ala	Ala	Leu 820	Asp	His	Lys	Ala	Asp 825	Ile	Ile	Gly	Leu 830	Ser	Gly	Leu
Ile	Thr 835	Pro	Ser	Leu	Asp	Glu	Met 840	Ile	Phe	Val	Ala	Lys 845	Glu	Met	Glu
Arg	Leu 850	Ala	Ile	Arg	Ile	Pro 855	Leu	Leu	Ile	Gly	Gly 860	Ala	Thr	Thr	Ser
Lys 865	Thr	His	Thr	Ala	Val 870	Lys	Ile	Ala	Pro	Arg 875	Tyr	Ser	Ala	Pro	Val 880
Xaa	His	Val	Leu	Asp 885	Ala	Ser	Lys	Ser	Val 890	Val	Val	Cys	Ser	Gln 895	Leu
Leu	Asp	Glu	Asn 900	Leu	Lys	Asp	Glu	Tyr 905	Phe	Glu	Glu	Ile 910	Met	Glu	Glu
Tyr	Glu	Asp 915	Ile	Arg	Gln	Xaa	Xaa 920	Tyr	Glu	Ser	Leu	Lys 925	Glu	Arg	Arg
Tyr	Leu 930	Pro	Leu	Ser	Gln	Ala 935	Arg	Lys	Ser	Gly	Phe 940	Gln	Met	Asp	Trp
Leu 945	Ser	Glu	Pro	His	Pro 950	Lys	Pro	Thr	Phe 955	Ile	Gly	Thr	Gln	Val 960	
Phe	Glu	Asp	Tyr	Asp	Leu	Gln	Lys	Leu	Val	Asp	Tyr	Ile	Asp	Trp	Lys

965 970 975
 Pro Phe Phe Asp Val Trp Gln Leu Arg Gly Lys Tyr Pro Asn Arg Gly
 980 985 990
 Phe Pro Lys Ile Phe Asn Asp Lys Thr Val Gly Gly Glu Ala Arg Lys
 995 1000 1005
 Val Tyr Asp Asp Ala His Asn Met Leu Asn Thr Leu Ile Ser Gln Lys
 1010 1015 1020
 Lys Leu Arg Ala Arg Gly Val Val Gly Phe Trp Pro Ala Gln Ser Ile
 1025 1030 1035 1040
 Gln Asp Asp Ile His Leu Tyr Ala Glu Ala Ala Val Pro Gln Ala Ala
 1045 1050 1055
 Glu Pro Ile Ala Thr Phe Tyr Gly Leu Arg Gln Gln Ala Glu Lys Asp
 1060 1065 1070
 Ser Ala Ser Thr Glu Pro Tyr Tyr Cys Leu Ser Asp Phe Ile Ala Pro
 1075 1080 1085
 Leu His Ser Gly Ile Arg Asp Tyr Leu Gly Leu Phe Ala Val Ala Cys
 1090 1095 1100
 Phe Gly Val Glu Glu Leu Ser Lys Ala Tyr Glu Asp Asp Gly Asp Asp
 1105 1110 1115 1120
 Tyr Ser Ser Ile Met Val Lys Ala Leu Gly Asp Arg Leu Ala Glu Ala
 1125 1130 1135
 Phe Ala Glu Glu Leu His Glu Arg Val Arg Arg Glu Leu Trp Ala Tyr
 1140 1145 1150
 Cys Gly Ser Glu Gln Leu Asp Val Ala Asp Leu Arg Arg Leu Arg Tyr
 1155 1160 1165
 Lys Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ser Gln Pro Asp His Thr
 1170 1175 1180
 Glu Lys Leu Thr Met Trp Arg Leu Ala Asp Ile Glu Gln Ser Thr Gly
 1185 1190 1195 1200
 Ile Arg Leu Thr Glu Ser Leu Ala Met Ala Pro Ala Ser Ala Val Ser
 1205 1210 1215
 Gly Leu Tyr Phe Ser Asn Leu Lys Ser Lys Tyr Phe Ala Val Gly Lys
 1220 1225 1230
 Ile Ser Lys Asp Gln Val Glu Asp Tyr Ala Leu Arg Lys Asn Ile Ser
 1235 1240 1245
 Val Ala Glu Val Glu Lys Trp Leu Gly Pro Ile Leu Gly Tyr Asp Thr
 1250 1255 1260
 Asp
 1265

<210> 75
 <211> 3856
 <212> DNA
 <213> Homo sapiens

<220>
 <221> variation
 <222> (1)...(3856)
 <223> nnn at positions 2640-2642 is either AAT or no
 nucleotides; n at position 2756 is either A or G;
 n at position 2758 is either C or G.

<400> 75
 atgtcacccg cgctccaaga cctgtcgcaa cccgaaggct tgaagaaaac cctgcgggat 60
 gagatcaatg ccattctgca gaagaggatt atgggtgctg atggagggat ggggaccatg 120
 atccagcggg agaagctaaa cgaagaacac ttccgaggtc aggaatttaa agatcatgcc 180
 aggccgctga aaggcaacaa tgacatttta agtataactc agcctgatgt catttaccaa 240
 atccataagg aatacttgct ggctggggca gatatcattg aaacaaatac ttttagcagc 300

actagtattg	cccaagctga	ctatggcctt	gaacacttgg	cctaccggat	gaacatgtgc	360
tctgcaggag	tgccagaaa	agctgccgag	gaggtaactc	tccagacagg	aattaagagg	420
tttgtggcag	gggctctggg	tccgactaat	aagacactct	ctgtgtcccc	atctgtggaa	480
aggccggatt	ataggaacat	cacatttgat	gagcttggtg	aagcatacca	agagcaggcc	540
aaaggacttc	tggatggcgg	ggttgatata	ttactcattg	aaactatttt	tgatactgcc	600
aatgccaaag	cagccttggt	tgcactccaa	aatctttttg	aggagaaata	tgctccccgg	660
cctatcttta	tttcagggac	gatcgttgat	aaaagtgggc	ggactctttc	cggacagaca	720
ggagagggat	ttgtcatcag	cgtgtctcat	ggagaaccac	tctgcattgg	attaaattgt	780
gctttgggtg	cagctgagat	gagacctttt	attgaaataa	ttggaaaatg	tacaacagcc	840
tatgtcctct	gttatcccaa	tgcaggtctt	cccaacacct	ttggtgacta	tgatgaaacg	900
ccttctatga	tgcccaagca	cctaaaggat	tttgctatgg	atggcttggt	caatatagtt	960
ggaggatgct	gtgggtcaac	accagatcat	atcagggaaa	ttgctgaagc	tgtgaaaaat	1020
tgtaagccta	gagttccacc	tgccactgct	tttgaaggac	atatgttact	gtctggtcta	1080
gagcccttca	ggattggacc	gtacaccaac	tttgtttaaca	ttggagagcg	ctgtaatggt	1140
gcaggatcaa	ggaagtttgc	taaactcatc	atggcaggaa	actatgaaga	agccttggtg	1200
gttgccaaaag	tgcagggtga	aatgggagcc	caggtgttgg	atgtcaacat	ggatgatggc	1260
atgtagatg	gtccaagtgc	aatgaccaga	ttttgcaact	taattgcttc	cgagccagac	1320
atcgcaaagg	tacctttgtg	catcgactcc	tccaattttg	ctgtgattga	agctgggtta	1380
aagtgtgcc	aagggaagtg	cattgtcaat	agcattagtc	tgaaggaagg	agaggacgac	1440
ttcttggaga	aggccaggaa	gattaaaaag	tatggagctg	ctatggtggt	catggctttt	1500
gatgaagaag	gacaggcaac	agaaacagac	acaaaaatca	gagtgtgcac	ccgggcctac	1560
catctgcttg	tgaaaaaact	gggctttaat	ccaaatgaca	ttatttttga	ccctaataatc	1620
ctaaccattg	ggactggaat	ggaggaacac	aacttgatat	ccattaattt	tatccatgca	1680
acaaaagtca	ttaaagaaac	attacctgga	gccagaataa	gtggaggtct	ttccaacttg	1740
tccttctcct	tccgaggaat	ggaagccatt	cgagaagcaa	tgcatggggg	tttcctttac	1800
catgcaatca	agtctggcat	ggacatggag	atagtgaatg	ctggaaacct	ccctgtgtat	1860
gatgatatac	ataaggaact	tctgcagctc	tgtgaagatc	tcctctggaa	taaagacctt	1920
gaggccactg	agaagctctt	acgttatgcc	cagactcaag	gcacaggagg	gaagaaagtc	1980
attcagactg	atgagtggag	aaatggccct	gtcgaagaac	gccttgagta	tgcccttggtg	2040
aagggcattg	aaaaacatat	tattgaggat	actgaggaag	ccaggttaaa	ccaaaaaaaaa	2100
tatccccgac	ctctcaatat	aattgaagga	cccctgatga	atggaatgaa	aattgttggt	2160
gatctttttg	gagctggaaa	aatgtttcta	cctcagggtta	taaagtcagc	ccgggttatg	2220
aagaaggctg	ttggccacct	tatccctttc	atggaaaaag	aaagagaaga	aaccagagtg	2280
cttaacggca	cagtagaaga	agaggaccct	taccagggca	ccatcgtgct	ggccactggt	2340
aaaggcgacg	tgcacgacat	aggcaagaac	atagttggag	tagtccttgg	ctgcaataat	2400
ttccgagtta	ttgatttagg	agtcatgact	ccatgtgata	agatactgaa	agctgctctt	2460
gaccacaaag	cagatataat	tggcctgtca	ggactcatca	ctccttcctt	ggatgaaatg	2520
atttttgttg	ccaaggaaat	ggagagatta	gctataagga	ttccattggt	gattggagga	2580
gcaaccactt	caaaaaacca	cacagcagtt	aaaatagctc	cgagatacac	tgcacctgtn	2640
nnccatgtcc	tggacgcgtc	caagagtgtg	gtggtgtggt	cccagctggt	agatgaaaat	2700
ctaaaggatg	aataacttga	ggaaatcatg	gaagaatatg	aagatattag	acaggncnat	2760
tatgagctct	tcaaggagag	gagatactta	cccttaagtc	aagccagaaa	aagtggtttc	2820
caaattggatt	ggctgtctga	acctcaccca	gtgaagccca	cgtttattgg	gacccaggtc	2880
tttgaagact	atgacctgca	gaagctggtg	gactacattg	actggaagcc	tttctttgat	2940
gtctggcagc	tccggggcaa	gtacccgaat	cgaggcttcc	ccaagatatt	taacgacaaa	3000
acagtaggtg	gagaggccag	gaaggtctac	gatgatgccc	acaatatgct	gaacacactg	3060
attagtcaaa	agaaactccg	ggcccggggt	gtggttgggt	tctggccagc	acagagtatc	3120
caagacgaca	ttcacctgta	cgcagaggct	gctgtgcccc	aggctgcaga	gcccatagcc	3180
actttctatg	ggttaaggca	acaggctgag	aaggactctg	ccagcacgga	gccatactac	3240
tgctctctcag	acttcatcgc	tcccttgcat	tctggcatcc	gtgactacct	gggctgtttt	3300
gcogttgcct	gctttggggt	agaagagctg	agcaaggcct	atgaggatga	tggtgacgac	3360
tacagcagca	tcatggtcaa	ggcgtggggg	gaccggctgg	cagaggcctt	tgcaagaagag	3420
ctccatgaaa	gagttcgccg	agaactgtgg	gcctactgtg	gcagtgagca	gctggacgtc	3480
gcagacctgc	gaaggttgcg	gtacaagggc	atccgcccgg	ctcctggcta	ccccagccag	3540
cccgaccaca	ccgagaagct	caccatgtgg	agactcgagc	acatcgagca	gtctacaggc	3600
attaggttaa	cagaatcatt	agcaatggca	cctgcttcag	cagtctcagg	cctctacttc	3660
tccaatttga	agtccaaata	ttttgctgtg	gggaagattt	ccaaggatca	ggttgaggat	3720
tatgcattga	ggaagaacat	atctgtggct	gaggttgaga	aatggcttgg	acccattttg	3780
ggatatgata	cagactaact	tttttttttt	tttttgcctt	ttttatcttg	atgatcctca	3840

aggaaataca acctag

3856

<210> 76

<211> 10

<212> DNA

<213> Homo sapiens

<400> 76

gacaacatgt

10